A diary study of information needs and document usage in the engineering domain

Peter J. Wild, Institute for Manufacturing, University of Cambridge, 17 Charles Babbage Road, Cambridge CB3 0FS, UK
Chris McMahon, Mansur Darlington and *Steve Culley,* Innovative Design and Manufacturing Research Centre, University of Bath, Bath BA2 7AY, UK
Shaofeng Liu, University of Plymouth, Business School, Hampton Street, Plymouth PL4 8AA, UK

Engineering design, despite its technical mechanical and mathematical foundations, is heavily informational. Many authors have used information as the focus of their research; in contrast, our study examined both information needs and document usage by engineers. After reviewing the design decisions behind the diary study; we report on the analysis of the data. We also used the data to generate new 'document use' scenarios, and 'proof of concept' test of a related software system. As with all research methods, diary studies have a number of strengths and weaknesses so we also provide some reflections on the use of diary studies as a method for examining the engineer's use of documents in the engineering domain.

Keywords: engineering design, information design, research methods, design studies

Engineering design practice, despite its technical mechanical and mathematical foundations, is also heavily informational (Court et al., 1996; Henderson, 1999; Lowe et al., 2004). For example, Lowe, et al's (2004) survey of engineers reported that 20 per cent of engineers' time was spent searching and absorbing information. The format, manifestation, and availability of this information are key variables in the success of design. Vajna (2005) summarised the situation thus, '*usually it is hard (if not impossible) to find the right documents, data or information at the right time. Therefore, it is difficult to finish the work quickly and to appropriate levels of quality* (p. 367).' However, previous studies in this area have focussed on information types (e.g. Court et al., 1996) rather than the concrete mechanism by which engineers record and distribute information, that is, the myriad forms of documents used in engineering and design. In engineering design, Roy et al (2004) investigated documentation procedures in aero-engine development, and observed that with paper-based documents there were issues with traceability and bottlenecks. Henderson (1999) investigated the importance

Corresponding author:
Peter J Wild
pw308@cam.ac.uk

www.elsevier.com/locate/destud
0142-694X $ - see front matter *Design Studies* **31** (2010) 46–73
doi:10.1016/j.destud.2009.06.002

of diagrams, and provided detailed evidence of 'mixed use' of both electronic and paper documents as common engineering practice. More generally there is wider interest in the properties of documents and their use in work (e.g. Levy, 1994; Adler et al., 1998; Zellweger et al., 2000; Sellen and Harper, 2002; McAlpine et al., 2006) as well as information use and management in engineering (Court et al., 1996; Hicks et al., 2006; McAlpine et al., 2006).

Current information search and retrieval methods provide the means to return candidate documents as potential sources of the information being sought. However, they usually leave the user with the task of exploring the document content unaided in the search for the information itself. It becomes necessary to find ways of identifying which, from a set of available, relevant documents, are useful, and which elements of information in a useful document will provide the most satisfactory information. It is necessary, too, to find means how to access those document elements effectively.

The research reported here was situated within a three-strand project. The first strand concerned how documents are structured, and how they can be decomposed according to different criteria (Darlington, 2005; Liu et al., 2006a; Liu et al., 2007). The second strand concerned Computational Methods & Mark-up Languages for documents (Liu et al., 2006b; Liu et al., 2007). The third strand, concerned, how designers use documents and navigate within them (Wild et al., 2005; Wild et al., 2006; Wild et al., 2009). The work reported here was carried out to understand information user need and document 'characteristics' and the impact of these on information provision to engineers. This paper is concerned with reporting on the use of a diary study to investigate information needs and their corresponding document access and usage.

The paper continues through an additional five sections. Section two provides an overview of diary studies as an empirical method and our diary design. Section three presents the results and analysis from our diary study. Section four presents a set of scenarios about information needs and document use derived from the data. Given the relative novelty of Diary Studies section five provides a set of reflections on the nature of diary studies and section six concludes the paper.

1 Diary studies as an empirical method

It is widely accepted that supporting designers through research demands empirical studies of how they carry out their tasks. However, Engineering design faces continuing difficulties of getting access to study real engineers in real-world settings (Subrahmanian et al., 2004). Engineer's time is often cited as one of the main reasons why subject participation is not forthcoming. Hence there is discussion about which research methods help to achieve these aims (Wild et al., 2007). A diary study was adopted, in part because it was considered less intrusive than experiments and protocol studies, and more

novel – to participants – than surveys and questionnaires. A second motivation is that the diary study would generate actual instances of document usage as reported by engineers, which could then be used to generate grounded scenarios of document usage in engineering domains.

1.1 Overview of diary studies

Diary studies are a form of research method concerned with *self-logging* of activity by participants (Zimmerman and Wider, 1977; Rieman, 1993; Pedgley, 2007). Examples vary in the literature, depending upon whether the researchers' predominant concern has been with generating predictions of behaviour (Higgins et al., 1985; Ericsson et al., 1993), understanding behaviour (Rieman, 1993; Havens and Schervish, 2001) or to inform the design of products (Adler et al., 1998; Czerwinski et al., 2004). The complexity of diary formats has also varied, from a simple format (i.e. date, time + free text) through to something more involved, which is akin to filling in multiple instantiations of a questionnaire. There has also been a characterisation of diary studies as being longitudinal (Pedgley, 2007) although this is not a defining feature within other work (Rieman, 1993; Czerwinski et al., 2004).

Diary studies have been used in a variety of domains and disciplines, for example in design of artefacts (Czerwinski et al., 2004) or treatment of pain occurrences (Ong et al., 2006). Their use in design-oriented disciplines has been greatest in Human-Computer Interaction (HCI) which reports over thirty diary studies undertaken on 'general' populations (i.e., non-engineers). In HCI they are used to gain insights into phenomena as well as serving as a source of design ideas (Czerwinski et al., 2004), and whilst some have been used to drive theory development (Adler et al., 1998), the use of diaries in the pursuance of definitive facts and figures is rare. In contrast, fewer diary studies of designers or engineers have been undertaken and these did not report on the use of documents (Ball et al., 1994; Dorst and Hendriks, 2001; Aurisicchio et al., 2003; MacGregor, 2004; Pedgley, 2007).

A weakness of diary studies is the lack of methodological guidance. Some pragmatic advice is given by Rieman (1993) and Corti (1993) but little is said about deeper design decisions. As a method, diary studies are methodologically and epistemologically ambiguous. The simple statement that 'we did a diary study' is not enough; exploration of the design decisions taken is also needed.

Crotty (1998) notes that '*in a very real sense, every piece of research is unique and calls for a unique methodology* (p. 13–14)'. This diary study was undertaken up with an exploratory mindset (Yin, 2003, p. 5–6). Whilst we report summaries of data, an additional concern within our work was to use the data to drive feature development of a toolset, the EDCMS (Liu et al., 2006a, 2008). This was done in part, by using data from the diary study to structure sample queries for proof of concept testing of the EDCMS; as well as

additional theoretical, empirical and implementation work that has been reported elsewhere (Darlington, 2005; Wild et al., 2005, 2006; Liu et al., 2006a, 2008). We cover more of the features of the EDCMS in Section 3.2. However, EDCMS can be viewed as embodying a basic principle that a document can be broken down or decomposed in multiple ways. Furthermore, these ways can be combined along with 'older' search methods (e.g., browsing and keyword search) to provide a method for providing relevant chunks of a document.

1.2 The diary study design

Within this section we outline a number of issues we tackled when designing the diary study.

1.2.1 Predictive or exploratory

The first issue was whether the diary would be used in a predictive or exploratory mode of empiricism. We could have developed a number of hypotheses and attempted to falsify them. Whilst the hypothetico-deductive approach can be used in an exploratory manner (Plutchik, 1974) it benefits from an established theory and set of observations from which to ground it. In contrast despite many valiant efforts, so little is known about information, knowledge, and document use in engineering design a set of meaningful hypotheses would be difficult. Thus, the diary study was undertaken with an exploratory mindset (Yin, 2003, p. 5–6), that is, what could we find out about information needs and document use in context, as a way of informing the design of a computer toolset as well as providing insights that are more general in nature (e.g., Document Usage Scenarios).

1.2.2 Scope: information, documents or both

Our first concern was with the scope of the diary study, and whether we focused solely on documents, or whether we considered some aspect of context through the notion of an information need (Wilson, 1981). Focussing solely on documents would have missed much of the important context or the activity for which a document is being used for. However, focussing solely on information needs alone would not be satisfactory since it would move us away from the focus of work which is document navigation and usage.

As a result of these deliberations, in our study, a diary entry had two parts. The first part is concerned with the 'information need', which asks questions about the activity being undertaken, the information being sought and the activities that were undertaken to do this. The second part is concerned with details about 'document access and use'. It requests detail about the document 'type', why the participant read the document, who 'owns' the document and what further actions where taken. Tables 2 and 3 list the questions asked in each part of a diary entry. Each information need could cross-reference one or more document access and usage forms. The overwhelming number of information needs correspond to a single document, although later in Section

3.2 we provide a more refined analysis of the ratios of documents to information needs.

1.2.3 Form of data gathered

The next design issue was the nature of the data being elicited. Our motivation with the diary study was to elicit data about instances of document access and use and their corresponding information need and activity. A variety of diary study instantiations exist gathering either qualitative or quantitative data or a mixture. At one extreme, the use of a date and time field with a blank sheet provides little structure for participants to respond. We suspected that, in the context of engineer's time constraints, this would result in too little information being filled in. Feedback from a pilot study suggested that more structure, rather than less, was needed.

At the 'quantitative' extreme, the user is asked to fill in a form that allows no deviation from set categories. This option allows the analysis of the data with extensive statistical measures, and may have the benefit that more entries are filled in. However, the problem with this is that little is known about information needs and documents in engineering to construct such a 'rigorous' elicitation structure. In the end a compromise was made between the two 'extremes'. Some questions (e.g. document type, document elements used, archives searched) had categories that participants tick, with an option to provide alternatives. Other questions (e.g. the information need, the activity, what follow-on actions occurred) were open in format and the participants were responsible for describing their own behaviour.

The suggested advantages to this compromise format were that it balanced the risks of meaningless quantitative data or too little relevant qualitative data being recorded, this would leave us in the position of interpreting results with little additional context for a quantitative entry (Labuschagne, 2003). Additionally, the data would be rich and sufficiently interesting to refine into a number of scenarios. It also allows the data to be analysed in the light of associated project research streams (i.e. document decomposition and computer tools). For example, the qualitative answers give insight into the nature of the searches being undertaken and the quantitative answers could help us to prioritise support in our prototype software tool for certain document types, search patterns and reading patterns.

1.2.4 Real-time entry?

The next issue is whether participants fill in the entries in real-time. Aurisicchio et al (2003) were strongly concerned that diary entries happened in this manner. In general, little is said by other diary studies as to whether diarists filled in diary entries in real-time. Our expectation from the pilot was that they would not do this all the time. Hence, we asked respondents to indicate when they

Table 1 When the diarists filled in their entries

Immediate	*1–10 min*	*10–60 min*	*>1 h*	*>1 day*
46 (33%)	33 (24%)	19 (13%)	23 (16%)	19 (13%)

had filled in the entry, the data in Table 1 shows that the largest category for filling in the diary was immediate.

1.3 Diary format and questions

The diary was available as a paper based, MS-Word or plain text based form. Diarists were evenly split in their preference for paper or electronic format, and there was no discernable difference in length of answers or attention to detail in either the electronic- or paper-based format. A detailed set of instructions were given as well as a number of sample entries. Tables 2 and 3 list the questions and indicate whether the participant had to enter a free text or check box answer.

2 Results and analysis

A total number of 142 'information needs' and 172 corresponding 'document and access use' entries were collected from 18 diarists over a 12-month period. All were involved in Engineering Design projects although we did not collect data on scale or duration of these projects. This section presents results and analysis of a number of representative questions from Tables 2 and 3.

2.1 An example entry

The content of Box 1 is taken directly from a diary entry by one our participants. Following anonymisation, the bulk of it is quoted verbatim from the diary entry; the further particulars are a summary of several responses, across the two parts of a diary entry (i.e., Information Need and Document Use and Access).

2.2 Data and interpretations of information needs

Within this section, our concern is to present data about the information needs that were collected in the diary and the interpretations and implication of this data. We cover a range of sample activities.

Table 2 Information elicited in the information need form

Question	*Information type*
Name	Free text
Date	Free text
Time	Free text
Time elapsed since activity	Checkbox
Activity being undertaken	Free text
Information been sought	Free text
New or old information need	Checkbox + optional free text
What has been done	Checkbox + optional free text
What document repository has been accessed	Checkbox + optional free text
What cues caused the document to be chosen	Free text
Has the information need been satisfied, and by which document(s)	Free text

Table 3 Information elicited in the document access and use form

Question	*Information type*
What 'type' is the document	Checkbox + optional free text
Why did you read the document	Checkbox + optional free text
What is the document format (electronic or paper based)	Checkbox + optional free text
Who does the document belong to	Checkbox + optional free text
Time taken to use the Document?	Free text
What elements of the document helped you find the information sought	Checkbox + optional free text
What sort of elements would help	Checkbox + optional free text
What did you do with the document	Checkbox + optional free text
Has the document satisfied the information need, and if so how	Free text
Did any other action occur	Free text

2.2.1 Sample activities

Box 2 lists a selection of the activities reported by participants.

We developed a provisional coding of these activities and have since refined this. The coding is based on recurring 'keywords' and synonyms in the

Box 1. An example diary entry

Activity

EMC Testing. A dispute has arisen over a number of comments in a customer Test Report issued 3 months ago. A search of the Test activity Logbook needed to be undertaken to check what was done and recorded to the time of testing. The Logbook has been archived (as paper)

Information sought

Procedure description of one particular test (there are a number available) from a particular specification document. The information is needed to clarify a point relating to data recorded in a customer test report already issued to the customer.

Need satisfied

The Logbook clearly referred to the technical question posed by the customer. As a result a letter was drafted to the customer to explain the point in question as the customer had misunderstood the comments in the test report and a more comprehensive explanation using reference material found in the test logbook was provided. At this time (until customer responds) it is not clear how successfully we have answer the customers questions.

Alternative action

At this time (until customer responds) it is not clear how successfully we have answered the customer's questions. It may be at a later date we need to respond again to clarify points with the customer.

Further particulars

The documents in question were local and paper based and an electronic file based on the workgroup area. The engineer read the documents to learn and to search for data and found the heading and abstract useful. No other action was taken to store or copy the documents.

Box 2. Sample participant activities

Preparing agenda for monthly project management meeting.
Technical information for engineers for reprogramming PLC and commissioning machine
Informing engineering colleagues (3 engineers) about modifications of inspection systems
Research into different disabilities to establish the varying levels of use of muscle groups for each types of physical disability.
Creating CAD model
Preparing project capital expenditure justification
External ground exchanges design slinky layout
Ground exchanges vertical layout
Business report market analysis
Update status report
Prepare for meeting
Reviewing Friday's team meeting minutes and preparing list of things to do today
Machinery design
Feasibility study of small scale wind turbine
Equipment design
Developing design
Handle design
CAD drawings

qualitative description of the activities. In some cases, multiple codes have been assigned to individual entries Table 4.

2.2.2 Sample of information needs

Box 3 lists a number of example information needs; these illustrate the information that is being sought by the participants. We planned to develop a coding scheme that would enable a consistent analysis of these, but found the task difficult due to the absence of a measure of abstractness in the literature. However, a sample of the information needs has been used to test-drive a number of hypothetical queries for the EDCMS toolset. This allows us to consider the applicable

Table 4 Coding of activities

Category	*Count*
Feasibility Study	02
Analysis	09
Design	04
Testing	02
Information gathering	41
Learning	09
Meeting	15
Informing	08
Planning	07
Documenting	40
Training	04

Box 3. Sample information needs

I heard about a clutch type differential which should be lighter and cheaper than the current used Torsen differential
I need to show and tell engineers what modifications are to be performed on the inspection machine the working principle of a refrigerator
Clean room standard
I need to know where the meeting is, with whom, and at what time (the meeting had been pre-arranged by another person I needed the information about it.
What are the varying levels of use of muscle groups for each types of physical disability?
I need to assess who was undertaking which tasks and evaluate whether they had achieved this.
Dimensions of equipment
Cost of waste for various products & various stages of manufacturing process
To find customer specs
Check details of minutes
General dimensions of main components and distances between components at the rear of the car.
General information on 2wd (two-wheel drive) motorcycles
How good/useful was the hydraulic 4-w-d system which was installed on BURT05?
What are the topics that have to be discussed and what has to be achieved by the end of the meeting?
What are the current gear ratio used in BURT05?
A general idea of what the different suspension systems are and which ones would be suitable for our car?
How do I select appropriate bearings
Background on how it works Different designs used Possible heat output
A spreadsheet report of data on usage of TWI's website
Expenditure and income for a collaborative project
Whether or not two EU Directives were common to one product or exclusive and independent of each other
Global costs of moulded plastic parts, broken down by geographical region
Past design alternatives suggested
Information on benefits of new product
Sketches from earlier today
Procedure description of one particular test (there are a number available) from a particular specification document
Anthropometric data, specifically bodyweight + hopefully the distribution of weight in the body
Composition of steel 4140
Was my recollection that we use BASIC ELECTRODES to keyword low-hydrogen covered electrodes right?
What is the typical microstructure of stainless steels 310 and 309?
Schedule information for facilities
Technical specification for facilities
Original map suggested revisions

Table 5 Example of new and old needs

New needs	*Old needs*
New project. Have to find out the general background of the project I am doing.	Needed updated project information to seek authorisation from director to the project
Enquiry came in by e-mail	"I've searched some information on the internet already
A one off query from a customer	A regular need
'Haven't used email address before it is the company that the person works for'	I took logbook with modification document to factory and showed engineers different aspects to be modified on machine
The arrival of a new version of a technical specification means it has to be closely checked for differences in the information we extract to set up our semiautomatic test systems	Found out previously about the system and about who had been in charge of the project so I could look for the associated report in the library
Heard mention of information by word of mouth. Now need additional information	Looking for documents of related titles to previous searches that appear to have more of an engineering focus
I used to remember this but have been away from this job for a few years so I am reluctant to trust my memory.	Use the library catalogue to find books that related to 'cleanroom'"
Various things requested at the meeting yesterday to be distributed to attendees	Preparation for a review meeting to discuss improvements to the spreadsheet report

decomposition schemes (Darlington, 2005, Liu et al., 2006b) that would support information needs. We cover this use of the diary data in more detail in Section 3.2.

2.2.3 New needs verses continuation of old needs

Diarists were asked to indicate whether the information need was new or old. From the 141 information needs, 82 were new, 62 old. Three were classed as both new and old, examination of the diarists answers and other entries suggested they understood the diary form, and where trying to explain a context and need that did not fit into either category However, we appreciate that for some this is a concept that would need refinement in future studies. Table 5 provides an example.

2.2.4 What was done to find the information

We asked diarists to report what action(s) they undertook to find the information. Table 6 lists the results.

Table 6 What activities were undertaken to find the information

Activity	*Count*	*Additional categories*	*Count*
Checked personal Memory	47	Other initial	67
Asked colleague	19	Search	25
Note to search later	8	Logbook	4
No answer	26	Document archive	11
		Final other	27

We see some reliance on personal memory as an activity used to find information. In 26 of the 141 data points, multiple activities were carried out; that is, they undertook more than one from: consulting memory; consulting colleagues; making a note to search later; or an additional activity.

2.2.5 Document repositories accessed

In seeking information from documents, people searched a number of different archives. We asked participants whether the archives belonged to them, their employer, or were external. We also asked whether the documents were paper or computer based. Despite the prevalence of organisational database and libraries a significant portion of documents are still being sought within personal file stores. In turn, a large portion of physical documents are being consulted and created. Overall, the results in Table 7 support the view that people still rely on paper based and personal archives for documents (Henderson, 1999; Sellen and Harper, 2002; McAlpine et al., 2006). This has implications for the form of empirical studies that are undertaken in the future into information and knowledge needs of engineers, an issue we return to in Section 5.

2.3 Data about document access and use

We now move onto to present summary data on the participants' document access and use. As can be seen in Table 3, the *document access and use* form concerned issues such as duration of the document use; how they read the document (after Adler et al., 1998) where the document was; how its manifestation (electronic/paper); and what they did with the document.

We would caution against arguing for or against this being fully representative of ***all*** document practices in engineering. In this study, we focus on the documents that are accessed and reported in relation to specific information needs. We have

Table 7 Document repositories accessed by the diarists

Source	*Count*
Personal archive	
Computer Based	30
Paper based	22
My employer's archive	
Workgroup	9
Library	15
Intranet	6
Not applicable	5
Colleague's archive	
Computer based	8
Paper based	3
External archive	
Library external	7
Web search	24
Online database	4

not been concerned with the documents that are pushed towards the diarist, such as traditional post, emails, attachments, faxes, reports, and papers for review, signing etc. Nor did we explicitly elicit data about those documents that were created. We did undertake an analysis of those documents that were being created by the diarists, and 45 are explicitly referred to as being created during the activities being reported. A number of the 'document access and use' entries referred to more than one document in the text. However, in the absence of the additional contextual data (e.g. manifestation, location, reading) we decided against splitting the *document access and use* forms into additional entries.

2.3.1 Ratios of documents to information needs

Each information need could cross-reference one or more document access and usage forms. The overwhelming number of information needs correspond to a single document. Figure 1 illustrates that the majority of the information needs had a one-to-one correspondence with a document. A number of information needs entailed (28 per cent) more than one document. Figure 1 shows the breakdown into number of documents consulted (1, 2, 3, 4, or 5) by the 141 Information Needs.

2.3.2 Duration of document use

We asked diarists to record how long they used each document. We then calculated the total time for each Information Need by totalling the amount of time spent with all the documents used in support of it. The results are listed in Figure 2. Categories for duration are under 2 min, 3–10 min, 11–60 min; 61–300 min; and finally, over 301 min.

Our interpretation of these durations is that often documents are used in a quick reference manner or merely to confirm that they are the correct document and passed on to colleagues or taken to meetings. In contrast, the time spent with the documents over two hours tended to involve more complex and drawn out activities. However, this does not of course guarantee that any full period is solely dedicated to a single document. Tables 7 and 8 provide a sample of the shorter and longer duration activities. Those longer document

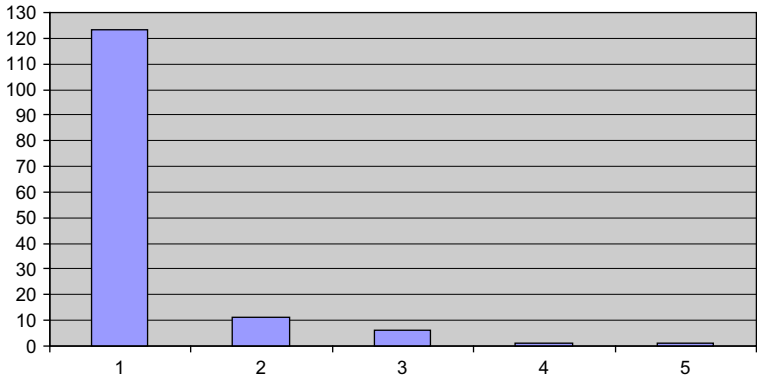

Figure 1 How many documents were accessed in relation to the Information needs

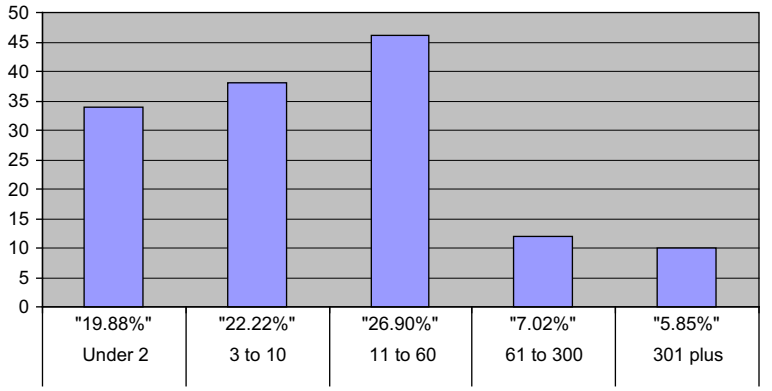

Figure 2 Duration of document use for each information need (n = 141)

usage periods (Table 9) were more intellectually engaging and involved activities than simply passing on of information (Table 8).

2.3.3 Document types accessed

We asked participants what was the 'type' of document they accessed. The results are in Table 10 – additional categories of document type are listed in the right hand section of the table. Related work (Wild et al., 2005; Wild et al., 2006) suggested that even a small engineering firm had over 250 different types of document, thus, a selection was piloted and included in the Document Access and Use form. These document types which are indicated on the left of Table 10.

2.3.4 Form of reading

In the *document access and use* form participants were asked how they read the document. This utilised an adapted taxonomy of work-related reading from Adler et al (1998) who developed a taxonomy of 8 categories of reading such as Skim read, Read to remind etc. When designing the diary format, we added two additional categories: Read to search for data; and Read to search for image or figure. The results had an additional 23 suggestions, from which we derived two additional categories of reading: *Read published material to remind*; and, at first glance, the intriguing, *Not read*. Each instance of *Not read* is qualified by the response to the follow-on action question 'need to read the book' and, in the second case, by the follow-on action 'sent email.' This indicates the usefulness of the additional contextual information in the diary entries. It also contrasts with the cruder analysis of tasks in other information related diary studies (e.g. Hyldegård, 2006).

Table 8 Example activities from shorter duration document uses

Activity	*Duration*
Checking that the keywords applied to an abstract by an outworker are correct, and making any changes or additions required.	01
Looking for email address to send update email	01
After meeting paperwork and emails	01
Preparing summary report	02
Circulating newsletter	03

Table 9 Activities and longer duration document use

Activity	*Duration*
Accessing engineering specification documents to gather information to construct a new format of Logbook (A logbook is the document in which a test engineer records the results of the testing undertaken)	150
Set up templates for gathering data on standard times for specifications	155
Accessing specification library to compile list of required tests from a number of the specifications	180
Analysing data flow in document production	270
Feasibility Study of small scale wind turbine	300
Feasibility study of wind turbine in Chew Magna UK on a large scale	420
Find out total energy consumption of Chew Magna	840
Creating CAD model	1050
Preparing project capital expenditure justification	1050
Background in heat pump design	1260
Find out Transco's opinion on grid connection of renewable energy sources	1470
Business report market analysis	1680
Ground exchanges vertical layout	2100
Machinery design	3990

No indication on the form was given whether to provide just one description or multiple reasons for reading the document. However, on over 40 occasions, participants gave multiple reasons for reading a document. Figure 3 illustrates the distribution.

To us, the design implications of this data are clear. Each one of these 'classes' of reading is a different reading task (Barnard, 1985; Adler et al., 1998). It would follow that they require different types of support needing, in turn, different access mechanisms, tools or representations. For example, the task of reading to support a discussion requires information about subject, pros, cons, existing arguments, etc. These could be viewed as decomposition elements (Darlington, 2005; Liu et al. 2006b, 2008). For example the category 'assess or critically review text' is likely to be 'read' in a more abstract manner,

Table 10 Types of document accessed

Document type	*Count*	*Analysis of 61 'others'*	
Drawing	08	Initial Others	61
Correspondence	12	Logbook	04
Journal	10	Website	14
Minutes	05	Presentation	03
Technical report	04	Reference Book	15
Management report	04	Final others	25
Standard	07		
Numerical	06		
Feasibility	02		
Supplier information	02		
Contract	01		
Multimedia	05		
Instruction manual	03		

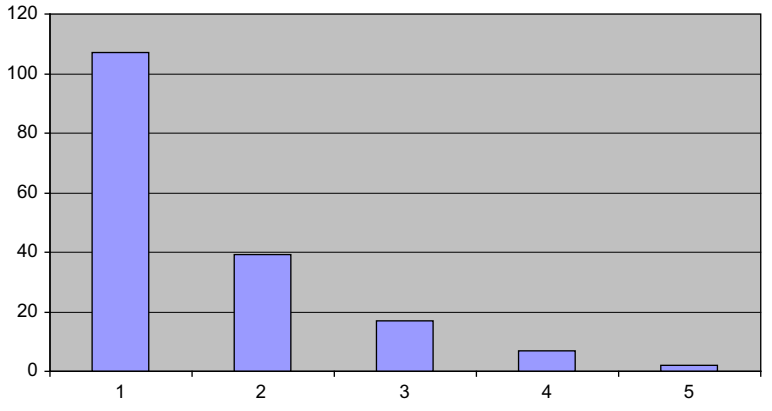

Figure 3 single or multiple types of document reading (n = 171)

revolving around propositional representation of the content, inferences about its meaning and evaluation of subsequent actions. In contrast, reading for 'image' or 'figure', is a searching and matching operation. This demands different configurations of cognitive resources (Barnard, 1985). Therefore the sorts of decomposition elements (Darlington, 2005; Liu et al., 2006b, 2008) needed for this task could be different than for other tasks.

2.3.5 Cues and document elements used

We also wanted to elicit information about the cues used to find the document and the elements of the document used, and the results are summarized in Table 12.

Table 11 reports on the cues used by diarists to find the actual document. As can be expected, these tend to be externally represented features, such as title, name, and date. The category 'pushed' refers to documents that were given to

Table 11 How the participants read the document[a]

Reading purpose	*Count*
Skim read	20
Read own text to remind	13
Read to search/answer questions	39
Read to learn	26
Read for cross-referencing	12
Read to assess, edit or critically review text	6
Read to support listening	0
Read to support discussion	12
Read to search for data[b]	48
Read to search for image or figure[b]	15
Initial Others	23
Not Read[c]	2
Read published text to remind[c]	3
Final Others	18

[a] Those options marked.
[b] Where added by the researchers.
[c] Derived from the data from analysis of the free text responses to 'other'. Those responses with more than two occurrences were promoted to categories in themselves.

Table 12 Cues used by diarists to find the document

Category	*Count*
Title	46
Pushed	8
Date	7
Name	4
Location	17
Memory	8
Keywords	1
Alphabetical list	6

the diarist, with typical responses being, 'it was e-mailed to me after I requested it earlier in the day', and 'sent by supplier'.

What is marked in the responses in the 'others' textbox is that other physical elements were requested but clearer demarcation of conceptual elements such as Introduction, Summary, etc. were not used or requested. Given the lower number of 'Read to assess, edit or critically review text (6)' in Table 11, this is not surprising (Table 13).

Diarists are using a wide variety of features when finding information within the document. We see several implications of this data for the design of document management systems. The data points to a diversity of potential clues in navigating to a document and within a specific document. This implies benefits for more support for 'regeneration' of documents with enhanced navigation structures, and reading structures. Examples of this include breaking strings of points e.g. 1) 2) 3) in a paragraph into bulleted points on separate lines; or the generation of 'tables of contents' for documents without altering other vital navigation information, such as page numbering and layout. This all points to enhanced coding and representation of the elements of the document, whether conceptual or physical. For example mark-up of data in images,

Table 13 The document elements used and wished for

Document element	*Used*	*Desired*
Abstract/summary	10	1
Keywords	15	4
Contents list,	14	2
Figure of table list	5	1
Table or figure caption	8	1
Section heading	24	3
Index	15	0
Search	11	5
NA	27	47
Other	23	6
Email subject	2	–
Bullet points	2	–
Alphabetical listing	6	–
Other final	13	–

Table 14 Copies or links made

Option	*Count*	*Option*	*Count*
Copied onto my computer	11	Other initial	45
Copied onto server	3	Already paper	3
Printed	18	Copied	21
Photocopied	2	Ready Ref	6
Bookmarked	3	Final others	15
None	72		
Added to citation list	4		

tables, etc. allows for more refined searching, for example by row and column, or by labels within an image. In addition, a reading profile could allow the EDCMS to reformat a document in response to a person's reading style or visual needs and constraints. Future fertile research could investigate the co-occurrence of document elements and reading patterns.

2.3.6 *Copy or link*

Our concern was whether diarists made copies or links of the documents with which they dealt. There were 158 responses, with 23 null entries, 138 had a single follow-on action, and there were 12 instances where there were 2 follow-on actions (Table 14).

2.3.7 *Additional actions*

We requested diarists to list additional actions and in 79 of the 172 data points additional actions were taken. Box 4 illustrates a sample of these additional activities.

These additional actions provide further evidence that documents are used within a context (of activities and needs), and that it is not always a case of use document, finish task. Documents are used within and embedded in long-lasting processes (McMahon et al., 2005). This includes documents that are considered traditional to design, but also documents concerned with collaboration, communication (see Hales, 1987).

3 *Using the data*

As well as presenting the results and our interpretation of their implications, we have developed a number of scenarios from the diary data that illustrate the wider ways that engineers use documents. These scenarios go beyond the simple consumer-producer information use characterisation that is the legacy of Shannon and Weaver's conceptualisation; to one of information being represented in documents used in a rich and complex fashion; from simple queries through to intensive multiple day activities with the same document.

Box 4. A sample of additional actions

E-mailed the agenda to all weekly meeting attendees
Kept copy and gave copy to engineer
Account will do some number crunching using IFS to provide me spending details of the whole project
Complete the question with the knowledge that I got
Make notes in my Logbook, print the document out.
Search some company/personal website that describe the standard in more detail. Though I cannot get the document, but at least I know the standard more.
Need to read the book
No. Because there should not have this product else the hospital will not ask us to do research on this topic.
I assessed whether they had completed this task and created a new source of information detailing this.
Edited incorrect details
Made a note in my logbook to find a supplier of differentials and email him for advice on which one to use
Invoice values sent to contracts department to produce invoice for company.
Asked member company if they require further information to be more specific in their materials definition.
Sketch design and create CAD model – still to do
At this time (until customer responds) it is not clear how successfully we have answered the customer's questions. It may be at a later date we need to responded again to clarify points with the customer.
The construction of a new Logbook format will require the process of reference to a number of different specification documents. This is an ongoing process at this work place at the present time.
The quoting/testing activity is monitored to check the success of the pricing policy. Actions taken then will depend on the success of the pricing policy.
Further consultation of other specification documents will be required in order to achieve a complete update of the training records
Search again in 1 month time, using slightly different keywords or another search engine.
Cross-referenced with XXX-DAU-38 and XXX-DAU-39

3.1 Grounded scenarios of information needs and document access

Scenarios are increasingly used in Design Research to express design issues, requirements etc. in a format that is accessible to the widest set of people working on the project, as well as industrially based stakeholders (Schmidt-Belz and Hovestadt, 1996; Go and Carroll, 2004; Lim and Sato, 2006). It is argued that scenarios are both sufficiently concrete and high level to communicate the essence of a situation of concern or interest without the need to learn detailed modelling techniques. The use of scenarios is a large and growing area and we

direct the reader to a number of sources if they wish to know the detail of specific methods (Schmidt-Belz and Hovestadt, 1996; Antón and Potts, 1998; Go and Carroll, 2004; Lim and Sato, 2006). Scenarios throughout various communities of practice rely on informality and concreteness to fuel discussion in relation to both design issues and theoretical issues. We put these scenarios as novel and interesting that go beyond situations encountered in related work (e.g., Adler et al., 1998; Aurisicchio et al., 2003; Roy et al., 2004). They should be seen as providing requirements for future document support software, not just ours.

To many these scenarios may not appear to characteristic of engineering design. However we cite precedent in Hales' (1987) study of activities undertaken by designers. Hales compared the activities undertaken against those suggested by Pahl and Beitz (1996). He found that Pahl and Beitz's categorisation of design phases covered around 47 per cent of the observed activities. Hence, these scenarios reflect both well-known activities (e.g. design) but also others concerning quick reference and preparation for meetings.

3.1.1 Focussed time with a document

Several of the diary entries reported that the diarist spent focussed time with a single document, from a period of a few hours through to several days. Hence this scenario concerns intensive use of a document for an extended time period, and these are periods of work that are task-driven rather than interrupt-driven (Miyata and Norman, 1986).

This scenario will be associated with certain reading patterns (Adler et al., 1998) and proofing activities, but overall reading patterns will cover the whole gamut of the categories found in Table 10.

It is likely that this scenario will entail multiple sub-versions of a document as alternative sections and designs are explored, detailed, evolved, folded back into the main document or abandoned. Design consideration entails support for multiple views, annotations, being able to explore alternative version of the document and then fold them back into the main version.

3.1.2 Revisiting a document on a regular basis

Many documents serve as logs of activity. These range from the humble engineer's logbook, (McAlpine et al., 2006) to documents that act as formal records of the progress of processes and tasks. A number of our diary entries referred to this pattern of behaviour with documents. A document was opened and additional information placed in it, often in a regular temporal cycle.

The key to understanding this scenario is to think in terms of the suspension and resumption of an activity and corresponding document(s) and context on a regular basis (Wild et al., 2004). Mechanisms such as audit trails, comments, etc. can all help the documenter/reader re-familiarise themselves with

the content and purpose of the document, more rapidly than an entire read through of the document. Auto-generation of abstract and summaries as well as highlighting of key elements of structure, such as the elements of the Logical decomposition scheme (Darlington, 2005; Liu et al., 2006b, 2008) would also support faster finding of the relevant sections.

3.1.3 Cloning a document

A common practice is to 'clone' an existing document. This goes beyond simple use of templates to a deeper replication of both the document, but also information concerning the rationale and design processes. This ties in with the notion of a variant design (Pahl and Beitz, 1996) where new designs are generated by varying a standard design.

This scenario will also be associated with certain reading patterns (Adler et al., 1998) and proofing activities. The designer will need to read the original documents and the specifications for the new design, comparing and contrasting them to assess what will be the same and what will be different.

Design support needs to consider how certain areas of the design are represented; those aspects most amenable to variation can be highlighted, whilst those perceived to be harder to change could be marked in a different manner.

3.1.4 Quick reference

Several diary entries can be seen to be quick reference in nature; either to quickly confirm a minor detail or to check that the correct document is being passed on to others. Producers of information have evolved a set of conventions, for structuring information sources such as dictionaries, thesauri. However, more support is needed for the structured information used within design, as this is not always as straightforwardly formatted and presented as a dictionary or thesaurus. For example, whilst information within a Parts Catalogues is standardised, searching for parts between catalogues can entail adapting to multiple formats, presentation styles and groupings of parts.

3.1.5 Passing on to others

Several diary entries can be seen to be an alteration of the quick reference scenario, whereby documents are checked to ensure they are the correct version to be passed on to someone else.

Support for file naming conventions, and fast recognition, especially amongst multiple versions of a document are needed to support this scenario.

3.1.6 Meeting preparation

Meetings are a common part of engineering work. Coordination, that is, the management of dependencies between activities within a process, is carried out at multiple levels of granularity.

However, meetings will generally take someone away from their regular work environment, and diarists often reported finding documents in advance of meetings so that they could. The documents needed within a meeting are a small but important part of the collaboration work carried out.

Computer based support of documents in and for meeting preparation, entails features such as copying them to temporary storage medium and generating print runs.

3.1.7 Creating procedural document

A number of the diarists were concerned with creating procedural documents, for example the activities reported as 'Accessing Engineering Specification documents to gather information to construct a new format of Logbook (A logbook is the document in which a test engineer records the results of the testing undertaken)' and 'Creating a test procedure document'.

Here, knowledge of the decomposition schemes (Darlington, 2005; Liu et al., 2006b; Liu et al., 2008) could benefit those developing such documents.

3.2 Proof of concept testing of queries in the EDCMS

The basic principle behind the EDCMS toolset is that a document can be broken down or decomposed in multiple ways and that these ways can be combined along with older search approaches (e.g., browsing and keyword search) to provide a method for providing relevant chunks of a document. Chunks can be relatively tangible, such as sections, headings etc., but also conceptual, such as introductions etc.

The principal exploratory decomposition schemes that have been used in the prototype system are described below. A more comprehensive list of elements and their definitions can be found in Liu et al. (2006b, 2008). A document reader with a **physical structure** view interprets a document as a composition of chapters, sections, paragraphs, headings, footnotes and so on. Exploration of the document can be achieved by navigation assisted by an implicit understanding of the 'grammar' of this structure. The **logical content** view considers the sort of elements in a document which by convention carry such labels as 'introduction', 'objectives', or 'conclusions'. So the document user with a logical content view interprets a document as a composition of content streams such as aims, introductions, objectives, and conclusions and so on. The **media type** *view* focuses on the types of information carrier in a document. In an engineering document, in addition to text, elements like diagrams, drawings, and graphs are familiar and important to a reader, and audio and video clips are becoming common elements. From a **technical view**, a document might provide a description of processes, products, customers and problems, etc. In meeting information needs, engineers may be interested in the descriptive information associated with conceptual entities like technique/method

description, schedule, and comparison. Documents can be of course interpreted from a **context view**, which provides information about the metadata of a document. For example, who wrote the document? Where was the document published? When was the document created and last modified? Having access to such information as subject, keyword, creation date and so on are important in document search and access.

The purpose of this analysis was to: a) to provide a basis for developing interesting question to demonstrate the power of the EDCMS; b) provide a set of questions to compare different search methods (e.g. paper-based, Google Desktop, full-index search, etc); and c) to assist in validating the content of the existing decompositions schemes (Liu et al., 2006b, 2008).

Table 15 contains questions extracted or paraphrased from the diary response analysis. The questions are those that seemed to provide a basis for validating the EDCMS approach and to illustrate its potential. The table contains three columns: a) the expression of the original information need; b) the need re-expressed to form a question in the engineering domain; and c) the resources that might be needed (including those not currently available in the Waypoint EDCMS implementation) to support query answering.

The following key can be used to interpret the acronyms in the third column. A convention based decomposition is based what is conventionally used, whilst post hoc decomposition refers to the actual structure Box 5.

4 *Reflections on the diary study*

4.1 *Useful features of diary studies*

A key feature of diary studies is that participants interpret their behaviour and represent it to researchers via a structured format. In general, this is done in real-time, as the activity is undertaken. Hence, in general, recall effects from filling in a survey at a later date are potentially minimised. The tendency to embellish or distort answers may also be less prevalent than questionnaire-based approaches. The problem of recounting how you were taught to do something, rather than how you actually do it may also be avoided.

The data generated in most cases is rich and generally the content of diary entries accords with other studies of reading, and information use in engineering (e.g. Adler et al., 1998; Lowe et al., 2004; Hicks et al., 2006; McAlpine et al., 2006).

4.2 *Methodological warnings*

However, we have methodological warnings for those wishing to adopt diary studies in the future. As we noted in Section 2, diaries have rarely been used in hypothetico-deductive modes of research. Our review of diary studies found

Table 15 Making EDCMS Queries from the Information Needs

Original Information Need	*Paraphrased query (in the engineering domain)*	*Resources required for response: decomposition (element: attribute value); keyword*
General information on 2wd (two-wheel drive) motorcycles	Find general information on two-wheel-drive motorcycles.	LCDS_CB (introduction) + LCDS_PH(introductory, overview, background) + keyword (two-wheel drive, motorcycle, single-track vehicle)
How good/useful was the hydraulic 4-w-d system which was installed on BURT05?	Find material that considers the strengths and weaknesses of the hydraulic 4-w-d system as used on the BURT05 car.	LCDS_PH (discussion, analysis, evaluation, lessons learned) + keyword (4-wheel-drive, BURT05)
What are the current gear ratio used in BURT05? A general idea of what the different suspension systems are and which ones would be suitable for our car?	What are the current gear ratio used in BURTO5? Find introductory material describing the different types of motor vehicle suspension system	PSDS (paragraph, table, heading) + keyword (gear ratio, BURT05) LCDS_CB (introduction) + LCDS_PH (introductory, overview) + TDDS (system description, system function, system comparison, Q&A) + keyword (motor vehicle suspension)
How do I select appropriate bearings Background on how it works Different designs used Possible heat output	Find descriptions of methods for the selection of bearings Find methods for calculating the heat loss in a diesel engine	TDDS (technique/method, procedure) + keyword (bearing selection) PSDS (equation) + keyword (heat loss, ICE) TDDS (technique/method, procedure analysis) + keyword (heat loss, inefficiency, diesel)
Whether or not two EU Directives were common to one product or exclusive and independent of each other	Find EC directives relating to the occupational health hazards of chrome electroplating solutions and of anodizing solutions	PSDS (title heading) + keyword (EC health protection directive) + keyword (chrome, anodizing)
Information on benefits of new product	Find information on the benefits of the proposed new design of city bus	LCDS_PH (aim, objective, critique, discussion) + TDDS (product description, product purpose, Q&A) + keyword (city bus, benefit, pros and cons)

only two diary studies that stated hypotheses – Ericsson et al (1993) and Higgins et al (1985). This latter work is most relevant to us. Higgins et al (1985) undertook a diary study of phone call behaviour and made comparisons between diary entries and the actual calls made as recorded on the actual exchange. Their hypothesis that people would under-report phone calls was confirmed. If these phenomena are carried over to diary studies in general we can assume that diarists will under-report phenomena.

Box 5. Key to interpreting acronyms in Table 15. Making EDCMS queries from the information needs

LCDS_CB = Logical content decomposition scheme (convention based)
LCDS_PH = Logical content decomposition scheme (post hoc)
PSDS = Physical structure decomposition scheme
TDDS = Technical description decomposition scheme

However, providing definitive evidence for or against our data on information needs and document usage will be hard. Despite the growth in the use of CAD and PDM tools, document use in engineering design retains a strong physical form (Henderson, 1999; Roy et al., 2004; Wild et al., 2006). With the availability of cheap scanners, printers, and digital pen technology, there is more of a blur between digital and paper documents. Saying definitively whether a document exists online or offline can be a difficult task, their fluidity is both debated (Levy, 1994) and exploited (Zellweger et al., 2000). For verification of computer-based document users of computer logging methods (e.g. Kort and de Poot, 2005) face design decisions concerning intrusion and interruption of activities. Does the tool simply sit in the background logging events and activities deemed relevant, or do they 'intervene' and attempt to gain purposive reasons behind the document use? Either way, how do we initially study people to figure out what is relevant for subsequent coding and analysis of data? Validation is a difficult area (Golafshani, 2003) the complexity of this diary study makes it a difficult to provide evidence that definitively confirms that data we have. Some would argue that this is a weakness, but a truer picture involves consideration of why we adopted a diary study. We concur with McGrath (1981) that empirical research attempts to balance conflicting desiderata; specifically, maximisation of context, control over behaviour, and generalisation. These issues are faced whenever we make a choice about adoption of empirical methods (McGrath, 1981). Our concern with this diary study was to maximise gathering of data about information needs and document usage ***in context***. In general, a case can be made for some of the figures; we have reported being broadly in line with other studies (e.g. Adler et al., 1998; Lowe et al., 2004). However, our own approach, whilst reporting on the figures and summaries of data, has been exploratory. We aimed to gain data about information needs and document usage, as there was little from which to work.

5 *Conclusions*

Diaries are a method of getting more fuzzy glimpses (McMahon, 2004) of engineers' work and information practices. As demonstrated the data generated, and the richness of diary entries, can be used directly to generate scenarios and query requirements for the development of software tools. However, from a research methods' perspective, diary studies are an interesting methodological beast. As well as having some of the advantages of other techniques, as a hybrid form of research technique, they also have disadvantages. This diary study has demonstrated that we can elicit rich and informative information about engineers' use of documents and their corresponding information needs, and use this to inform design. However, those seeking full verification of the data will be disappointed.

We have demonstrated that the data from a diary study can be used to inform the design of engineering support tools. In the first case, we have developed a number of scenarios from the diary data that we can use to

consider how engineers use documents. These scenarios go beyond the simple consumer-producer characterisation that is the legacy of Shannon and Weaver's conceptualisation, to one of information being represented in documents used in a rich and complex fashion, from simple queries through to intensive multiple day activities with the same document. In the second case, we have demonstrated that the information needs can be used to provide proof of concept tests of the functionality of EDCMS and the decomposition schemes.

Key aspects of the data identified are that diarists: despite the presence of the workgroup and internet based resources — made considerable use of their personal archives; made considerable use of paper-based documents; undertook a range of activities ranging from short information finding and recording activities to activities that are longer in duration.

In turn, we have three key pieces of advice for those wishing to use diary studies in the future.

Firstly, if using diary studies predictively, use them in an established domain — do not expect to build simple comparison hypotheses (e.g. expert vs novice). Consider the complexity of the activity, and how complex activities are difficult to verify or falsify.

Secondly, if using a diary study in an exploratory mode of investigation, expect surprises, rich data, but a feeling of ambivalence regarding what the data means.

Finally, pursue multiple paths, never rely on one study site or organisation, and acknowledge that even one entry can provide insights into the phenomena being investigated.

Acknowledgements

The co-operation of various diarists from the following organisations is greatly appreciated: Rolls Royce, Vitacress, TWI, TUVPS, and Converteam. The research reported in this paper was undertaken at the Innovative Design Manufacturing Research Centre in Bath University funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant GR/R67507/01.

References

Adler, A, Gujar, A, Harrison, B L, O'Hara, K and Sellen, A (1998) A diary study of work-related reading: design implications for digital reading devices, *Proceedings of ACM CHI 98 Conference on Human Factors in Computing Systems* Vol 1 pp 241–248

Antón, A I and Potts, C (1998) A representational framework for scenarios of system use, *Requirements Engineering* Vol 3 No 3/4 pp 219–241
Aurisicchio, M, Langdon, P M, Ahmed, S, and Wallace, K M (2003) Investigating knowledge searches in aerospace design. *ICED03, 14th International Conference on Engineering Design*, Stockholm, Sweden, pp 293–294
Ball, L J, Evans, J and Dennis, I (1994) Cognitive processes in engineering design, *Ergonomics* Vol 37 No 11 pp 1753–1786
Barnard, P J (1985) Interacting cognitive subsystems: a psycholinguistic approach to short-term memory in **A Ellis** (ed) *Progress in the Psychology of Language*, Lawrence Erlbaum and Associates, London Vol 2 pp 197–258
Corti, L (1993) Using diaries in social research, *Social Research Update* Vol 2
Court, A W, Culley, S J and McMahon, C A (1996) Information access diagrams: a technique for analysing the usage of design information, *Journal of Engineering Design* Vol 7 pp 55–75
Crotty, M (1998) *The foundations of social research meaning and perspective in the research process* Sage, London
Czerwinski, M, Horvitz, E and Wilhite, S (2004) A diary study of task switching and interruptions in **A F Monk and D Olsen** (eds) *CHI 2004*, ACM Press, Vienna pp 175–182
Darlington, M J. (2005) Document Decomposition for Engineering Information Access and Support: Issues and Related Research. Report, ICID Internal Report No. 04/05, University of Bath, Bath.
Dorst, K and Hendriks, D (2001) The role of the design context in **P Lloyd and H Christiaans** (eds) *Practice and in Design Methodology* Designing in Context, Delft University of Technology, 18–20th December
Ericsson, K A, Krampe, R and Tesch-Romer, C (1993) The role of deliberate practice, *Psychological Review* Vol 100 No 3 pp 363–406
Go, K and Carroll, J M (2004) Scenario-Based Task Analysis in **D Diaper and N A Stanton** (eds) *The Handbook of Task Analysis for HCI*, Lawrence Erlbaum Associates, Mahwah, NJ pp 117–134
Golafshani, N (2003) Understanding reliability and validity in qualitative research, *The Qualitative Report* Vol 8 No 4 pp 597–606
Hales, C (1987) Analysis of the engineering design process in an industrial context, unpublished Ph.D. Thesis, Cambridge University, Cambridge.
Havens, J J and Schervish, P G (2001) The methods and metrics of the Boston area diary study, *Nonprofit and Voluntary Sector Quarterly* Vol 30 No 3 pp 527–550
Henderson, K (1999) *On Line and On Paper* MIT Press, Cambridge MA
Hicks, B J, Culley, S J and McMahon, C A (2006) A study of issues relating to information management across engineering SMEs, *International Journal of Information Management* Vol 26 No 4 pp 267–289
Higgins, C A, McClean, R J and Conrath, D W (1985) The accuracy and biases of diary communication data, *Social Networks* Vol 7 pp 173–187
Hyldegård, J (2006) Using diaries in group based information behavior research: a methodological study in *Proceedings of the 1st International Conference on Information Interaction in Context*, ACM, Copenhagen, Denmark pp 153–161
Kort J. and de Poot, H. (2005) Usage Analysis: Combining Logging and Qualitative Methods. *Conference on Human Factors in Computing Systems: CHI '05 extended abstracts on Human factors in computing systems*, Portland, Oregon, USA, April 2-7, 2005. pp 2121–2122.
Labuschagne, A (2003) Qualitative research - Airy fairy or fundamental? *The Qualitative Report* Vol 8 No 1

Levy, D M (1994) Fixed or fluid? document stability and new media in *Proceedings of the 1994 ACM European Conference on Hypermedia Technology*, ACM Press, Edinburgh pp 24–31

Lim, Y -K and Sato, K (2006) Describing multiple aspects of use situation: applications of Design Information Framework (DIF) to scenario development, *Design Studies* Vol 27 No 1 pp 57–76

Liu, F, McMahon, C, Darlington, M J, Culley, S and Wild, P J (2006a) An exploration of advanced computational technologies to facilitate retrieval of engineering document fragments, *Advanced Engineering Informatics* Vol 20 No 4 pp 401–413

Liu, S, McMahon, C A, Darlington, M J, Culley, S J and Wild, P J (2006b) An approach for document fragment retrieval and its formatting issue in engineering information management in **V Kumar**, **M L Gavrilova and O Gervasi** (eds) *International Conference on Computational Science and its Applications*, Springer-Verlag, Glasgow, Berlin pp 279–287

Liu, S, McMahon, C A, Darlington, M J, Culley, S J and Wild, P J (2007) EDCMS: a content management system for engineering documents, *International Journal of Automation and Computing* Vol 4 No 1 pp 56–70

Liu, S, McMahon, C A, Darlington, M J, Culley, S J and Wild, P J (2008) An automatic mark-up approach for structured document retrieval in engineering design, *International Journal of Advanced Manufacturing Technology* Vol 38 pp 418–425

Lowe, A, McMahon, C and Culley, S (2004) Information access, storage and use by engineering designers, part 1, *Engineering Designer* Vol 30 No 2 pp 30–32

MacGregor, S (2004) Using the case study method to study distributed design in **E Subrahmanian**, **R Sriram**, **P Herder**, **H Christiaans and R Schneider** (eds) *Workshop Proceedings: The Role of Empirical Studies in Understanding and Supporting Engineering Design*, Delft University Press pp 18–29

McAlpine, H, Hicks, B J, Huet, G and Culley, S J (2006) An Investigation into the use and content of the engineer's logbook, *Design Studies* Vol 27 No 4 pp 481–504

McGrath, J (1981) Dilemmatics: the study of research choices and dilemmas, *American Behavioral Scientist* Vol 25 No 2 pp 179–210

McMahon, C (2004) Empirical design research: fuzzy glimpses of practice in **E Subrahmanian**, **R Sriram**, **P Herder**, **H Christiaans and R Schneider** (eds) *Workshop Proceedings: The Role of Empirical Studies in Understanding and Supporting Engineering Design*, Delft University Press pp 103–113

McMahon, C, Giess, M and Culley, S (2005) Information management for through life product support: the curation of digital engineering data, *International Journal of Product Lifecycle Management* Vol 1 No 1 pp 26–42

Miyata, Y and Norman, D A (1986) Psychological issues in support of multiple activities in **D A Norman and S Draper** (eds) *User Centred systems Design: New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum and Associates, Hillsdale, NJ pp 265–284

Ong, B N, Dunn, K M and Croft, P R (2006) 'Since you're asking… ': free text commentaries in an epidemiological study of low back pain consulters in primary care, *Quality and Quantity* Vol 40 No 4

Pahl, G and Beitz, W (1996) *Engineering Design – A Systematic Approach* (2nd edn) Springer-Verlag, London

Pedgley, O (2007) Capturing and analysing own design activity, *Design Studies* Vol 28 No 5 pp 463–483

Plutchik, R (1974) *Foundations of Experimental Research* (2nd edn) Harper and Row, New York

Rieman, J. (1993) The Diary Study: A Workplace-Oriented Research Tool to Guide Laboratory Efforts. *Proceedings of ACM INTERCHI'93 Conference on Human Factors in Computing Systems*, pp 321–326.
Roy, R, Makri, C and Kritsilis, D (2004) Documenting technical specifications during the conceptualisation stages of aeroengine product development in **D Marjanovic** (ed) *8th International Design Conference* pp 897–902 Dubrovnik, Croatia, May 18–21
Schmidt-Belz, B and Hovestadt, L (1996) Scenario of an integrated design support for architects, *Design Studies* Vol 17 No 4 pp 489–509
Sellen, A and Harper, R H R (2002) *The Myth of the Paperless Office* MIT Press, Cambridge MA
Subrahmanian, E, Sriram, R, Herder, P, Christiaans, H and Schneider, R (2004) *The Role of Empirical Studies in Understanding and Supporting Engineering Design: Workshop Proceedings* Delft University Press, Delft
Vajna, S (2005) Workflow for Design in **P J Clarkson and C M Eckert** (eds) *Design Process Improvement: A Review of Current Practice* pp 366–385 London
Wild, P J, Culley, S, McMahon, C, Darlington, M J and Liu, S (2006) Towards a method for profiling engineering documentation in **D Marjanovic** (ed) *Design 2006* pp 1309–1318 Dubrovnik, May 15–18th
Wild, P J, Culley, S, McMahon, C A, Darlington, M J and Liu, S (2005) Starting to audit documents in the engineering domain in **P Amaldi**, **S P Gill**, **B Fields and W Wong** (eds) *In-Use, In-Situ: Extending Field Research Methods* London, 27–28th October
Wild, P J, Giess, M D and McMahon, C A (2009) Describing engineering documents with facetted approaches: observations and reflections. *Journal of Documentation*, Vol 65 No 3 pp 420–445
Wild, P J, Johnson, P and Johnson, H (2004) Towards a composite model for multitasking in **P Palanque**, **P Salvik and M Winckler** (eds) *TAMODIA 2004*, ACM Press, Prague November 15–16
Wild, P J, McMahon, C A and Culley, S J (2007) Introduction to special issue on empirical approaches in design information and knowledge, *Journal of Design Research* Vol 6 No 3 pp 289–294
Wilson, T D (1981) On user studies and information needs, *Journal of Librarianship* Vol 37 No 1 pp 3–15
Yin, R K (2003) *Case Study Research: Design and Methods* (3rd edn) Sage Publications, London
Zellweger, P T, Harkness Regli, S, Mackinlay, J D and Chang, B -W (2000) The impact of fluid documents on reading and browsing: an observational study in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, The Hague, The Netherlands pp 249–256
Zimmerman, D H and Wider, L D (1977) Diary-interview method, *Urban Life* Vol 5 No 4 pp 479–498